Audio Engineering Society

Convention e-Brief 557

Presented at the 147th Convention
2019 October 16–19, New York, USA



# Perceptually Motivated Hearing Loss Simulation for Audio Mixing Reference

Angeliki Mourgela[1], Trevor Agus[2], and Joshua D. Reiss[3]

[1] *Queen Mary University of London, London, E1 4NS, United Kingdom*

[2] *Queen's University Belfast, Belfast, BT7 1NN, United Kingdom*

[3] *Queen Mary University of London, London, E1 4NS, United Kingdom*

Correspondence should be addressed to Angeliki Mourgela (a.mourgela@qmul.ac.uk)

**ABSTRACT**

This paper proposes the development of a hearing loss simulation for use in audio mix referencing, designed according to psychoacoustics and audiology research findings. The simulation proposed in this paper aims to reproduce four perceptual aspects of hearing loss; threshold elevation, loss of dynamic range, reduced frequency and temporal resolution, while providing an audio input/output functionality.

## 1 Introduction

Hearing loss, can have a great impact on the quality of life of an individual, affecting daily activities negatively, including communication as well as entertainment [5]. It is estimated that approximately 466 million people worldwide suffer from some form of hearing loss (HL) and this is projected to increase to over 900 million by the year 2050 [29]. The rapid growth in affected population has motivated audiology and psychoacoustics researchers towards investigating the physiological and perceptual aspects of hearing loss, with the aim of gaining a deeper insight, and providing effective solutions on improving the affected population's quality of life.

Simulating hearing loss can be a valuable tool towards exploring both the perceptual and physiological aspects of the affected auditory functions [8]. It also provides a useful evaluation tool for the signal processing required for effective HL compensation [9]. Previous HL simulation approaches include those based on the use of pre-processed stimuli (pure tones, masking noise etc.) that aim to trigger physiological responses resembling those of a listener with HL, as well as simulations that produce processed audio (either real time or offline) that demonstrate the experience and perceptual characteristics of hearing loss. A list of previous simulation approaches and their characteristics can be found in Table 1.

Additionally, audio mix referencing techniques are a common practice in the studio. They provide engineers with an effective way to test their mixes and prepare them for different playback conditions, thus ensuring compatibility and retaining mix quality and clarity regardless of reproduction systems. By introducing the HL simulation at the mixing stage, the engineer can evaluate the effects of the simulation processing on their mixes and perform the necessary adjustments for compensation [1,9,22,23,28].

This paper proposes a HL simulation, with audio input/output functionality, that aims to reproduce aspects of sensorineural hearing loss using real-time audio signal processing. The simulation reproduces the effects of mild to moderate loss, while offering customizability of severity of the hearing loss in different frequency regions. The simulation is designed to be used as a master bus stereo plugin, thus reproducing the effects of combined loss in both ears. It also allows for user input of audiogram threshold values, as well as severity adjustments in supra-threshold effects. Key perceptual aspects of sensorineural hearing loss that are modeled include:

- Threshold elevations, usually greater for higher frequencies [16]
- Loss of the compressive nonlinearity functions of the basilar membrane [18]
- Reduced frequency selectivity, caused by broadening of the auditory filters [13][24]
- Reduced temporal resolution, due to neural asynchrony mostly observed in auditory aging [11][19]

The simulation is designed to be used as a referencing tool in audio production, thus promoting a link between audio engineering, audiology and psychoacoustics. Benefits of using this development include promotion of awareness of the perceptual differences occurring with hearing loss and facilitation of the evaluation process for compensation mixing techniques and adjustments for HL audiences [1,9,22,23,28].

## 2 Technical Description

The proposed HL simulation is designed and coded in MATLAB and is comprised of four digital signal processing stages, connected in series so that each output is the next processor's input. A signal flow chart of the system is given in Fig. 1. The four stages of the simulation are as follows:

- Audiogram Matching
- Dynamic Range Processor (DRP)
- Spectral Smearing Processor
- Temporal Jitter Processor

### *2.1.1 Processing Stages*

At the first stage of the simulation, a frequency-sampling based FIR filter is used to replicate the threshold elevations outlined in the audiogram. The attenuated audio signal is then passed into a crossover filter, where it is separated into 4 frequency bands. Separation frequencies of the crossover are adjustable, so that the suprathreshold effects apply to the most affected frequency ranges. After separation, the first band, containing the lower frequencies, bypasses the dynamic and spectral processing and is reintroduced at the temporal jittering stage, before the bands are summed again to form the output signal. Each of the remaining three bands is then separately sent to the next signal processor.

The next stage of the simulation is a dynamic range processor (DRP). The DRP is designed so that its input/output level characteristics closely resemble the curve characterizing loss of compression in the basilar membrane [18], a phenomenon commonly observed in hearing loss and attributed to outer hair cell damage (Fig. 2). More specifically, a threshold of expansion ($T_{exp}$) is set to correspond to the audiometric thresholds found at the audiogram. An additional upper threshold of total recruitment ($T_{rec}$) is set to correspond to the level where complete loudness recruitment occurs, so that the signal level would sound equally loud to a HL listener as it would to a normal hearing (NH) listener. Any signals above $T_{exp}$ and up to $T_{rec}$, present steep rapid upwards expansion, which represents abnormal loudness growth and loss of compression. For signal levels equal to and greater than $T_{rec}$, the DRP will present a linear behavior.

| Reference | Method | Aspect Simulated |
|---|---|---|
| [10],[17] | Compressive Gammachirp Filter | Loss of Compression Threshold Elevation |
| [3],[4],[6],[7] | Multiband Dynamic Expansion | Loss of Compression |
| [3],[12],[16] | Gaussian Bell Convolution | Spectral Smearing |
| [2] | Low-Passed Noise Multiplication | Spectral Smearing |
| [3],[19] | Sample Shift | Temporal Jitter |
| [27],[31] | Automatic Gain Control | Loss of Compression Threshold Elevation |
| [6],[20] | Additive Threshold Noise | Threshold Elevation |

Table 1: Previous Simulation Approaches

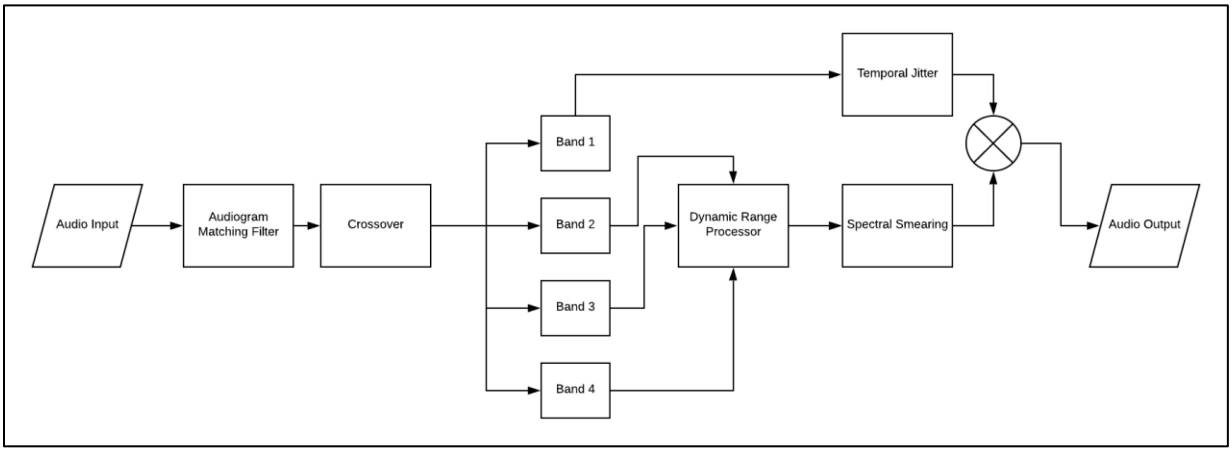


Figure 1: Simulation Signal Flow Chart

After exiting the DRP, each of the three processed sub-bands is passed through a separate spectral processor that introduces spectral smearing, representing the phenomenon of reduced frequency selectivity attributed to the widening of the auditory filters. Normal and widened auditory filter shapes are presented in Fig. 3 [24]. To simulate the effect of spectral smearing, a time domain multiplication of each of the three sub-band audio signals is performed with low passed noise filtered at different cutoff frequencies for each of the sub-bands [2].

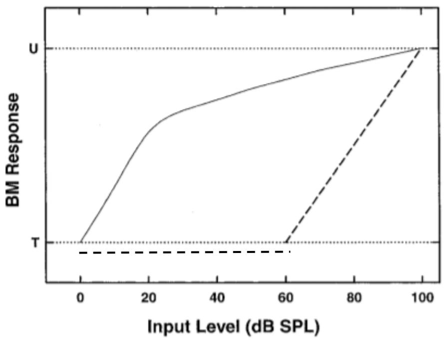


Figure 2: Input-Output level, characterising the Basilar Membrane response in normal (solid line) versus hearing impaired (dashed line) listeners. (adapted from: Oxenham and Bacon, 2003)

More specifically, the processor initially generates and filters the low-passed noise signals. The generated white noise is then convolved with a set of FIR low pass filters, with different cutoff frequencies, producing three different filtered noise signals. Each of the noise signals is multiplied in the time domain with its corresponding DRP audio output to produce the smeared signals.

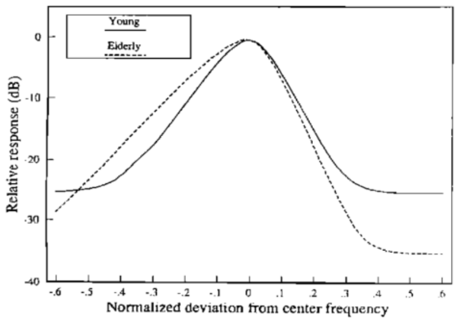


Figure 3: Auditory filter shapes. The solid line shows filter shapes for young listeners; the dashed line shows filters for elderly subjects. (source: Sommers and Humes, 1993)

At the last stage of the simulation, temporal jitter is introduced, only on the first low frequency band of the audio signal. This stage aims to reproduce the effects of neural asynchrony and loss of temporal resolution in the auditory system, an aspect of hearing loss linked to auditory aging, that mainly affects the lower frequencies [11,19].

To reproduce the effect of temporal jitter, the audio signal passes through a chorus-based processor, with a maximum delay value of 0.25ms, which has been found to significantly decrease intelligibility [19]. This technique introduces small and random delays within the signal.

Finally, the 4 bands are summed together to produce the output of the simulation.

## 3 Results & Discussion

### *3.1.1 Results*

The following section outlines the results of implementing the digital signal processors on test signals, as well as an evaluation of their effectiveness.

Processed audio samples of the output of the simulation can be accessed at: https://code.soundsoftware.ac.uk/projects/perceptually_motivated_hearing_loss_simulation/files

*Audiogram – matching filter*

The spectrum presented in Fig. 4 shows the frequency response of a white noise signal before and after the attenuation applied from the audiogram matching filter. The affected frequencies in the figure are those above 500 Hz, presenting a gradual moderate loss, as depicted in the audiogram presented in Fig. 5.

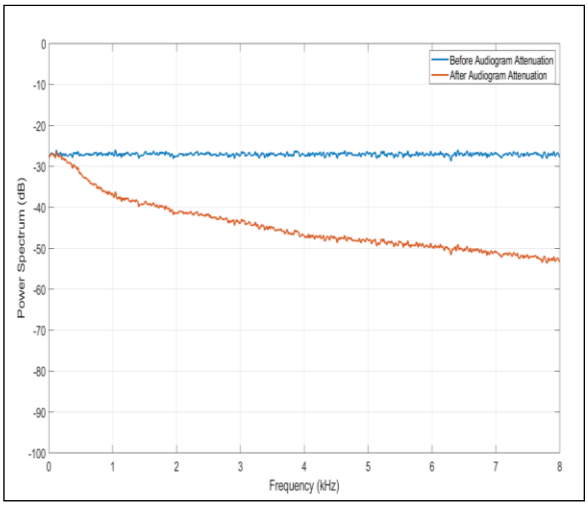


Figure 4: Frequency spectrum of a white noise signal before and after the FIR audiogram matching filter.

*Dynamic Range Processor*

The waveforms plotted in Fig. 6, present the input and output plot of the DRP. The input signal used here, is a speech signal (*source: freesound.org*), after it has been attenuated by the audiogram matching filter. The blue waveform in the graph presents the speech signal before the DRP while the red waveform presents the signal after the DRP.

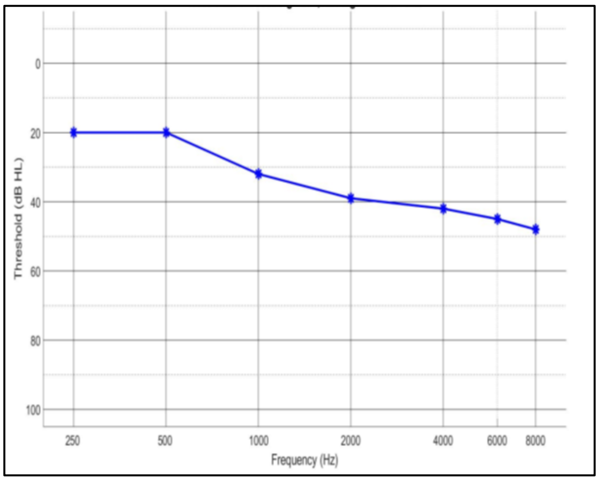


Figure 5: Audiogram representing moderate loss in both ears (generated in MATLAB)

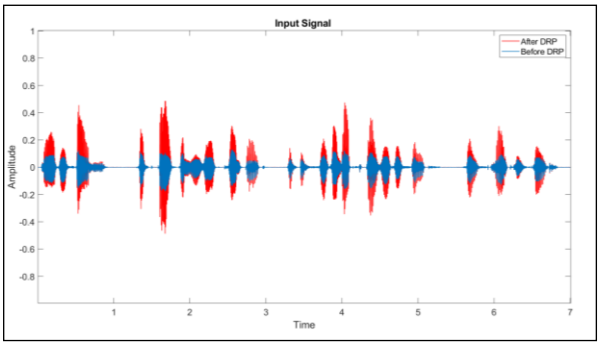


Figure 6: Input (blue) and output (red) waveform of speech signal, before and after the DRP.

*Spectral Smearing Processor*

The graphs shown in Fig. 7 present the input and output frequency spectrums of the spectral smearing processor. The input signal used here is a 1 kHz pure tone and the output is its time domain multiplication with a lowpass filtered white noise signal (low pass cut-off at 150 Hz).

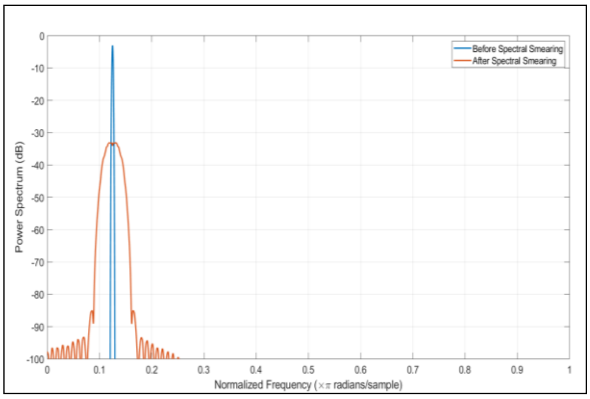


Figure 7: Input/output spectrum of 1kHz pure, before and after the spectral smearing processor.

The temporal smearing processing technique implemented in this paper, is based on chorus techniques adapted from [21],[26],[30]. The graphs of a triangular pulse before and after the temporal jitter processor are presented in Fig. 8.

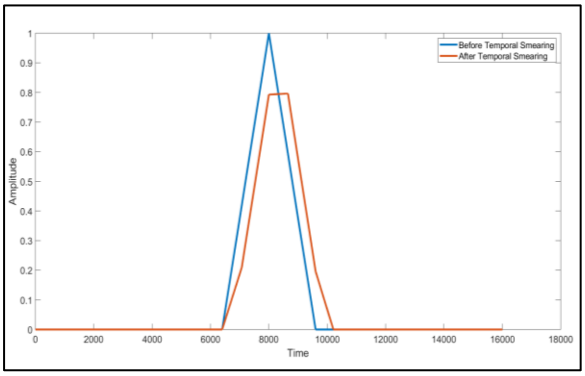


Figure 8: Input/output spectrum of triangular pulse signal before and after the temporal smearing processor.

### *3.1.2 Discussion*

The initial results of the simulation present a first approximation of the perceptual aspects of hearing loss, as documented in relevant studies and previous approaches. Moreover, the produced audio is in line with previous simulations [4][12]. However, it can be observed that the system presents certain limitations in its effectiveness. Such limitations include inaccuracies such as, a more rapid slope of attenuation at the audiogram filter stage, compared to that given by the audiogram, as well as audible artefacts at the output of the DRP, due to fast attack and release times.

The results of the simulation are very important to this study, as they provide the basis upon which the final development will be built. These results help identify the strongest and weakest points of the simulation, as well as set the direction for future improvements.

## 4 Conclusions

The development of this simulation demonstrates how some key perceptual aspects characterizing HL can be approximated by using audio signal processing techniques. Simulating hearing loss comes with certain limitations in its accuracy [9], whereas evaluating its effectiveness and validity mainly relies on psychophysical measurements and listening experiments. Limitations arise due to the complexity of the auditory perception, as well as the inability of a listener with HL to validate that their loss's perceptual characteristics resemble those reproduced by the simulation. Furthermore, hearing loss simulations that use digital signal processing (DSP) are bound to approximations, internal errors or processing limits. This simulation also presents certain limitations in its fidelity. These occur due to DSP constrains, as well as the difficulty in producing audio outputs for all aspects of HL, due to the fact that the data available in the related studies mainly correspond to values characterizing internal mechanisms of the cochlea (e.g. stapes velocity, basilar membrane displacement), or nerve firing rates.

Future developments include design improvements for better approximation and elimination of unwanted audible artefacts, as well as the transfer of the design to real-time processing. Moreover, the simulation will be remodelled into a plug-in format, in order to offer Digital Audio Workstation (DAW) compatibility, thus facilitating its embodiment into the mixing process. These developments will allow for the simulation to be both used and evaluated by engineers in practice, as well as motivate the further exploration of its use in producing enhanced audio mixes for audiences with HL.

## 5 Acknowledgements

This work is supported by the EPSRC and the BBC.

## 6 References

1. Agus, T. and Corrigan, C. (2018). *Adapting Audio Mixes for Hearing Impairments*. 3rd Workshop on Intelligent Music Production.

2. Boothroyd, A., et al. (1996). *Effects of spectral smearing on phoneme and word recognition*. Journal of the Acoustical Society of America, 100(3), pp.1807-1818.

3. Braida, L. and Lum, D. (2019). DSP *Implementation of a Real-Time Hearing Loss Simulator Based on Dynamic Expansion*. Modeling Sensorineural Hearing Loss, pp.113-129.

4. Cuevas-Rodriguez, M., et al. (2018). *The 3D Tune-In Toolkit – 3D audio spatialiser, hearing loss and hearing aid simulations.* IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE).

5. Dalton, D., et al. (2003). *The Impact of Hearing Loss on Quality of Life in Older Adults*. The Gerontologist, 43(5), pp.661-668.

6. Desloge, J., et al. (2010). *Speech reception by listeners with real and simulated hearing impairment: Effects of continuous and interrupted noise*. Journal of the Acoustical Society of America, 128(1), pp.342-359.

7. Duchnowski, P. and Zurek, P. (1989). *Simulation of sensorineural hearing loss*. The Journal of the Acoustical Society of America, 85(S1), pp.S26-S26.

8. Gagné, J. and Erber, N. (1987). *Simulation of Sensorineural Hearing Impairment. Ear and Hearing,* 8(4), pp.232-243.

9. Graf, I. (1997). *Simulation of the Effects of Sensorineural Hearing Loss.*

10. Grimault, N., et al. (2016). *Real-time simulation of hearing impairment: Application to speech-in noise intelligibility*. Journal of the Acoustical Society of America, 140(4), pp.3438-3438.

11. Gordon-Salant, S. and Fitzgibbons, P. (1993). *Temporal Factors and Speech Recognition Performance in Young and Elderly Listeners*. Journal of Speech, Language, and Hearing Research, 36(6), pp.1276-1285.

12. Huckvale, M. (2011). HearLoss - *Hearing Loss Demonstrator. [online] Phon.ucl.ac.uk.* Available at: https://www.phon.ucl.ac.uk/resource/hearloss/ [Accessed 30 Aug. 2019].

13. Léger, A., et al. (2012). *Effects of spectral smearing on the identification of speech in noise filtered into low- and mid-frequency regions*. Journal of the Acoustical Society of America, 131(5), pp.4114-4123.

14. McAllister, H., et al. (1994). *Audiogram matching using frequency sampling filters*. Proc. 16th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society.

15. Milekhina, O., et al. (2017). *Compressive Nonlinearity in the Auditory System: Manifestation in the Action of Complex Sound Signals*. Biology Bulletin, 44(6), pp.603-609.

16. Moore, B. and Glasberg, B. (1993). *Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech*. Journal of the Acoustical Society of America, 94(4), pp.2050-2062.

17. Nagae, M., et al. (2014*). Hearing impairment simulator based on compressive gammachirp filter.* Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA).

18. Oxenham, A. and Bacon, S. (2003). *Cochlear Compression: Perceptual Measures and Implications for Normal and Impaired Hearing*. Ear and Hearing, 24(5), pp.352-366.

19. Pichora-Fuller, et al. (2007). *Temporal jitter disrupts speech intelligibility: A simulation of auditory aging.* Hearing Research, 223(1-2), pp.114-121.

20. Plack, C., et al. (2008). *Estimates of compression at low and high frequencies using masking additivity in normal and impaired ears*. Journal of the Acoustical Society of America, 123(6), pp.4321-4330.

21. Reiss, J. and McPherson, A. (2015). *"Audio Effects: Theory, Implementation and Application"*. Boca Raton, Fla.: CRC Press.

22. Shirley, B., Meadows, M., Malak, F., Woodcock, J. and Tidball, A. (2017). *Personalized Object-Based Audio for Hearing Impaired TV Viewers.* Journal of the Audio Engineering Society, 65(4), pp.293-303.

23. Shirley, B. and Oldfield, R. (2015). *Clean Audio for TV broadcast: An Object-Based Approach for Hearing-Impaired Viewers*. Journal of the Audio Engineering Society, 63(4), pp.245-256.

24. Sommers, M. and Humes, L. (1993). *Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners).* Journal of the Acoustical Society of America, 93(5), pp.2903-2914.

25. Stables R., et al. (2019), *Intelligent Music Production, Routledge*, Taylor and Francis.

26. Tarr, E. (2018). *Hack audio: An Introduction to Computer Programming and Digital Signal Processing in MATLAB*. 1st ed. London: Taylor & Francis Ltd.

27. Villchur, E. (1974). *The Effect of Recruitment on Speech Perception—A Simulation for Normal Listeners.* Journal of the Acoustical Society of America, 55(2), pp.450-450

28. Ward, L. and Shirley, B. (2019). *Personalization in Object-based Audio for Accessibility: A Review of Advancements for Hearing Impaired Listeners.* Journal of the Audio Engineering Society, 67(7/8), pp.584-597.

29. Who Int. (2019). *Deafness and hearing loss.* [online] Available at: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

30. Zölzer, U. (2011). *DAFX. 2nd ed*. John Wiley & Sons, Ltd.

31. Zurek, P. and Desloge, J. (2004). *Immersive simulation of hearing loss and auditory prostheses*. Journal of the Acoustical Society of America, 115(5), pp.2602-2602.